IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received July 21, 2020, accepted August 7, 2020, date of publication August 17, 2020, date of current version August 28, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3016956

Improved PV Forecasts for Capacity Firming

CHANAKA KEERTHISINGHE[1,2], (Member, IEEE), EVAN MICKELSON[1], DANIEL S. KIRSCHEN[1], (Fellow, IEEE), NATHAN SHIH[1], (Student Member, IEEE), AND SCOTT GIBSON[3], (Member, IEEE)

[1]Department of Electrical and Computer Engineering, University of Washington, Seattle, WA 98195, USA
[2]Systems Integration Testbed, Washington Clean Energy Testbeds, Seattle, WA 98105, USA
[3]Snohomish County Public Utility District, Everett, WA 98201, USA

Corresponding author: Chanaka Keerthisinghe (chanakak@ieee.org)

ABSTRACT Some balancing authorities give owners of medium to large photovoltaic (PV) generation plants a choice between firming the production of their plants using battery energy storage or paying the balancing authority for the cost that these imbalances impose on the system. If the owner of a PV plant decides to do capacity firming, the net production of the PV plant and the battery must match a forecast value. A more accurate forecast of the PV production reduces the energy throughput of the battery and hence its degradation. This article compares capacity firming using persistence forecasts with predictions based on long short-term memory recurrent neural networks (LSTM-RNN), encoder-decoder LSTM-RNN and multi-layer perceptrons. This article also proposes to use the type-of-day, such as sunny, cloudy etc, which can be generated by clustering historical PV generation data according to the total daily PV generation, as a feature of the PV forecasting model. Results based on the Snohomish County Public Utility District's Arlington Microgrid show that the machine learning techniques perform significantly better than the persistence method in forecasting PV generation. In particular, encoder-decoder LSTM-RNN would reduce the yearly battery energy throughput by 29% and the number of battery cycles with a greater than 10% depth-of-discharge (DoD) by 51%. Including the day-type as a feature in PV forecasting reduces the battery energy throughput by 5.3% and the number of cycles with a DoD larger than 10% by 5.9%.

INDEX TERMS Capacity firming, solar smoothing, PV forecasting, encoder-decoder long short-term memory networks, recurrent neural networks, multi-layer perceptrons, battery energy storage, battery degradation, sequence2sequence time-series prediction.

I. INTRODUCTION

Providing the resources needed to balance the increasing amount of naturally variable and uncertain generation from solar and wind can be very costly for the balancing authorities (BA) that are responsible for maintaining the balance between load and generation within their territory. Some have therefore given the owners of these resources a choice between paying a balancing fee and firming up the output of their plants using their own resources.

This work is motivated by the capacity firming requirements of medium to large photovoltaic (PV) generation plants in the balancing area of the Bonneville Power Administration (BPA) and presents results based on the 500 kWac PV system of the Arlington Microgrid owned by Snohomish County Public Utility District (Snohomish PUD) [1]. Snohomish PUD can either pay a fee for capacity firming to BPA or compensate deviations between the actual and forecast PV generation using the 1MW/1.385MWh battery energy storage system (BESS) of the microgrid.

Capacity firming is based on a one-hour-ahead target PV generation determined considering the PV forecast, the maximum allowed ramp rate of the PV profile and the battery state-of-charge (SoC) that should be achieved at the next time-step. The BESS is charged when the PV generation is above the target value or discharged when the PV generation is below this value. In most cases the BESS has enough power and energy capacity to achieve this goal. However, the more inaccurate the forecast is, the more the battery has to compensate with deeper cycles. Since deeper cycles cause more battery degradation [2], [3], it may be more economical in the long run to pay the fee for capacity firming to BPA rather than firming the capacity using the battery. Since more accurate forecasts extend the life of the battery,

The associate editor coordinating the review of this manuscript and approving it for publication was Ning Kang.

This work is licensed under a Creative Commons Attribution 4.0 License. For more information, see https://creativecommons.org/licenses/by/4.0/

a careful analysis of the benefits of improved accuracy is thus required.

This article compares state-of-the-art PV forecasting techniques, such as long short-term memory recurrent neural networks (LSTM-RNN), encoder-decoder LSTM-RNN, multi-layer perceptrons (MLP), and the persistence method suggested by BPA. The paper also proposes to use the type-of-day as a feature in the PV forecasting model. The type-of-day is generated by clustering historical PV generation data according to the total daily PV output. The cluster with the highest PV generation is classified as sunny while the cluster with the least PV generation is classified as cloudy. These forecasting methods are compared using not only the root-mean-square error (RMSE), mean-absolute error (MAE) and mean-bias error (MBE), but also the resulting number and depth of the battery cycles.

The remainder of this article is structured as follows: Section II explains capacity firming; Section III reviews the existing PV forecasting techniques and discusses the machine learning based forecasting techniques; Sections IV and V describe the model, while Section VI presents and discusses the simulation results. Section VII concludes the paper.

II. CAPACITY FIRMING

This section explains the incentive scheme for capacity firming (Section II-A), formulates this problem mathematically (Section II-B) and describes the PV forecasting technique proposed by BPA (Section II-C).

A. INCENTIVE SCHEME FOR PV CAPACITY FIRMING

BPA, the balancing authority for the Snohomish PUD service region has decided to charge $5400 per year as the cost of capacity firming for the 500 kWac PV system of the Arlington Microgrid. Snohomish PUD could save this cost by ensuring that the actual PV production matches their one-hour-ahead prediction. If Snohomish PUD chooses this option, BPA has proposed to use the persistence approach for PV forecasting described in Section II-C.

B. FORMULATION OF THE CAPACITY FIRMING PROBLEM

The one-hour-ahead PV generation can be firmed using a BESS. The battery is charged when the PV generation is above the target value and discharged when the PV generation is below this value. Fig. 1. shows the energy flow diagram for this capacity firming application of the BESS.

The decision to do capacity firming or not depends on a cost/benefit analysis. While Snohomish PUD could avoid paying this fee to BPA if they did capacity firming, the extra battery cycling required would increase battery degradation and carry a long term cost. More accurate PV forecasts would require fewer battery cycles, cause less battery degradation costs and hence improve the cost/benefit ratio from capacity firming.

The capacity firming problem can be formalized as follows:

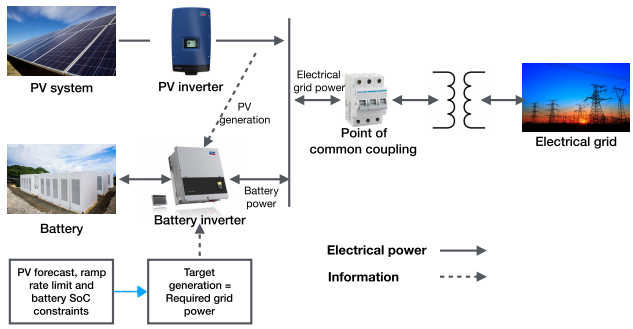


FIGURE 1. Electrical power and information flows in a PV-battery microgrid. The battery power is calculated using (1).

- *A sequence of time-steps*, $\mathcal{K} = \{1 \ldots k \ldots K\}$, where K and k represent the total number of time-steps and a particular time-step in a day, respectively. We choose 30 minute intervals because this is the duration that BPA uses to assess capacity firming.
- *A set of variables* to define the energy flows in the system. The battery charge/discharge rate p_k^{b}, is the difference between the actual PV generation $p_k^{\mathrm{pv,actual}}$ and the target output p_k^{t}:

$$p_k^{\mathrm{b}} = p_k^{\mathrm{t}} - p_k^{\mathrm{pv,actual}} \tag{1}$$

  where a positive value for p_k^{b} means the battery is discharging and a negative value that is charging.
- *The target output* p_k^{t} is estimated considering the PV forecast, the maximum allowed ramp rate of the PV profile and the desired battery SoC at the next time step. The target PV generation is calculated as follows:

$$p_k^{\mathrm{t}} = p_k^{\mathrm{pv,forecast}} + 2C\left(B_k^{\mathrm{SoC}} - B^{\mathrm{SoC},50\%}\right) \tag{2}$$

  where B_k^{SoC} is the battery SoC at time-step k in kWh, $B^{\mathrm{SoC},50\%}$ is the battery 50% SoC in kWh and C determines how close to the 50% level the SoC should remain. A higher value of C keeps the SoC closer to 50%. The factor "2" is needed to convert energy into power because we use 30 minute intervals. The BESS system is assumed to have enough power and energy capacities to correct the expected range of deviations in PV generation. Finally, the target PV generation submitted to BPA cannot violate the ramp rate limits:

$$\left|p_k^{\mathrm{t}} - p_{k-1}^{\mathrm{t}}\right| < \gamma_{\mathrm{r}} \tag{3}$$

  where γ_{r} is the ramp rate limit.

The battery SoC, $B_k^{\mathrm{SoC}} \in [s^{\mathrm{SoC,min}}, s_k^{\mathrm{SoC,max}}]$, evolves according to:

$$B_{k+1}^{\mathrm{SoC}} = B_k^{\mathrm{SoC}} - 0.5\eta_{\mathrm{b}} p^{\mathrm{b}} \tag{4}$$

where η_{b} is the battery efficiency.

The number of battery cycles that a lithium-ion battery can undergo before its capacity falls below an acceptable threshold depends on the depth-of-discharge (DoD) of

these cycles. The number of battery cycles and their magnitudes are calculated using the rainflow cycle counting algorithm and these numbers are then used to quantify the battery degradation.

The net present value (NPV) of capacity firming is calculated by subtracting the present value of the battery replacement cost after the lifetime of the project from the present value of all the revenue over the lifetime of the battery. The battery replacement cost is the expected degraded battery capacity over the lifetime multiplied by the battery replacement cost per kWh.

C. PERSISTENCE APPROACH TO FORECASTING

Because of its simplicity, BPA has suggested to Snohomish PUD to use the persistence forecasting technique for one-hour-ahead PV forecasting. For example, under 30/30 persistence forecasting, the net generation for the 2:00 PM to 2:30 PM interval is calculated by taking the average of the generation output from 1:00 PM to 1:30 PM. Similarly, the schedule for the 3:00 PM to 3:30 PM interval is calculated by taking the average of the generation output from 2:00 PM to 2:30 PM.

III. PV FORECASTING USING MACHINE LEARNING

This section first presents a brief review of existing PV forecasting techniques and then the state-of-the-art techniques that have been shown to provide the best PV forecasts.

A. BRIEF REVIEW OF EXISTING SOLUTION TECHNIQUES

A range of techniques have been proposed to optimize capacity firming [4]–[10] but these approaches did not include state-of-the-art machine learning techniques discussed in this article.

The underlying PV forecasting problem is a sequence-to-sequence (seq2seq) time series prediction because the PV generation has to be predicted two time-steps ahead. PV forecasting can be categorized into physical models and data-driven models. Physical models use numerical weather prediction, which shows good performance for forecast horizons from several hours up to six days [11], [12]. Data-driven models can be further divided into statistical and machine learning models. Statistical models include auto-regressive integrated moving average, auto-regressive moving average, coupled auto-regressive and dynamic system, Lasso, and Markov models. Machine learning models include support vector machine, feed-forward neural networks, and RNN such as LSTM networks [13]–[28]. These approaches can be further subdivided according to the type of input features that are used to train the model. A forecasting model that uses only a target time-series as an input feature (solar irradiance in this case) is referred to as a nonlinear auto-regressive (NAR) model. On the other hand, if a model uses additional exogenous inputs, such as temperature and humidity, it is referred to as a nonlinear auto-regressive with exogenous inputs (NARX) model. According to [13], [18], a vector output LSTM-RNN performs the best at forecasting day-ahead PV generation.

Section III-B describes feedforward neural networks, also known as multi-layer perceptrons, while Section III-C is devoted to LSTM-RNN and encoder-decoder LSTM-RNN. The encoder-decoder LSTM-RNN was first proposed and applied to speech recognition problems [29], [30] in order to effectively solve seq2seq time-series problems with multiple outputs. Our encoder-decoder LSTM-RNN implementation is based on [18].

B. MULTI-LAYER PERCEPTRONS

In general, MLPs are feedforward neural networks with multiple layers of perceptrons. A perceptron is a single neuron model that has weighted input signals and produces an output signal using an activation function.

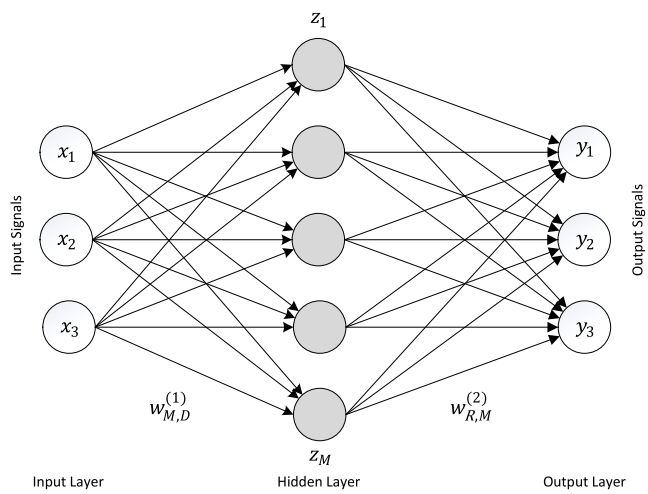


FIGURE 2. Feed-forward neural network. The circles represent the nodes of the input, hidden, and output layers. The arrows show the direction of forward propagation.

An MLP consists of at least three layers: an input layer, one or more hidden layers and an output layer, as depicted in Fig. 2. The input, hidden and output nodes are defined as $\mathbf{x} = [x_1 \ldots x_d \ldots x_D]$, $\mathbf{z} = [z_1 \ldots z_m \ldots z_M]$ and $\mathbf{y} = [y_1 \ldots y_r \ldots y_R]$, respectively. Here the total number of input, hidden and output nodes are D, M and R, while d, m and r denote a particular node. The output value for y_r is calculated using:

$$y_r(\mathbf{x}, \mathbf{w}) = \sigma\left(\sum_{m=1}^{M} w^{(2)}_{rm} \mathcal{H}\left(\sum_{d=1}^{D} w^{(1)}_{md} x_d + b^{(1)}_m\right) + b^{(2)}_r\right), \quad (5)$$

where σ and $\mathcal{H}$ are the activation functions for output and hidden layers, respectively. The biases $b^{(1)}_m$ and $b^{(2)}_r$ are for the hidden node m and the output node r, respectively. The weight matrices of hidden-input and output-hidden are $w^{(1)}_{md}$ and $w^{(2)}_{rm}$, respectively.

MLPs utilize a supervised learning technique called back propagation for training and the training data needs to be arranged in a 2-dimensional (2D) matrix of samples

and features. Therefore, the only way to use MLPs for time-series prediction is to use the time sequence data as a feature (more details in Section IV). The other option is to use a RNN, which is explained in the next section.

C. LONG SHORT-TERM MEMORY RECURRENT NEURAL NETWORKS

LSTM networks are a special type of RNN that are capable of learning long-term dependencies. Fig. 3 illustrates the RNN and Fig. 4 the LSTM cells [13].

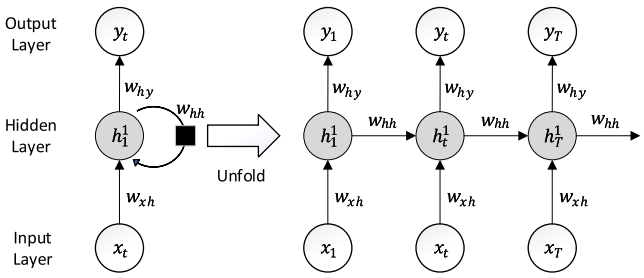


FIGURE 3. Recurrent neural network. The circles represent the input, hidden and output layers and the weights are above the lines. The arrows show the direction of forward propagation.

A RNN has a feedback connection where the output depends on the current input to the network, and the previous inputs, outputs, and/or hidden states of the network, as depicted in Fig. 3. Given an input sequence $\mathbf{x} = (x^1 \ldots x^t \ldots x^T)$, RNNs compute the hidden vector $\mathbf{h} = (h^1 \ldots h^t \ldots h^T)$, and output vector sequence $\mathbf{y} = (y^1 \ldots y^t \ldots y^T)$ by iterating (6) and implementing the calculation shown in (7).

$$h_t = \mathcal{H}\left(W_{xh} \cdot x_t + W_{hh} \cdot h_{t-1} + b_h\right), \tag{6}$$

$$y_t = W_{hy} \cdot h_t + b_y, \tag{7}$$

where $\mathcal{H}$ is the activation function of the hidden layer. W_{xh}, W_{hh} and W_{hy} are the weight matrices of the input-hidden, hidden-hidden, and hidden-output connections, respectively. The hidden and output bias vectors are b_h and b_y.

The error gradients of RNNs that are trained using back-propagation through time can accumulate during an update and result in very large gradients. These gradients result in large updates to the network weights, and in turn, an unstable network. At an extreme, the weights can become so large as to overflow and result in non-computable values. LSTM networks have been developed to overcome the exploding back-propagated gradients of RNNs by providing explicit memory to the network. LSTM units are building units for layers of RNNs. A typical LSTM unit consists of an input gate, forget gate, output gate, and a cell unit.

The operations of the LSTM unit can be described as follows. The most important component is the cell state c_t, which serves as a memory and remembers values over an arbitrary time interval. The input gate i_t, forget gate f_t, and output gate o_t control the flow of information into and out of the cell and has the same size as the hidden vector $\mathbf{h}$.

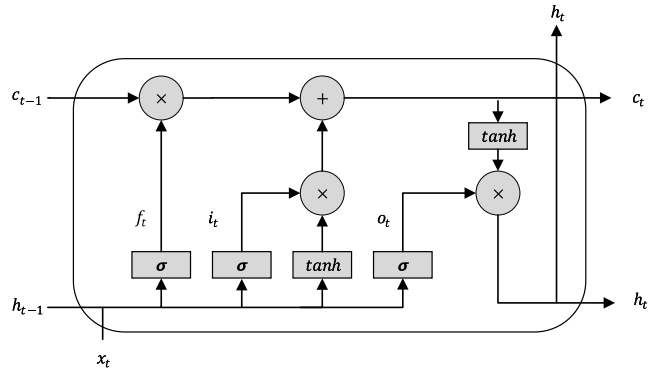


FIGURE 4. Long short-term memory cell.

The forget gate f_t outputs a 0 or 1 for each number in cell state to decide what information we want to put out from the previous cell state c_{t-1}, according to:

$$f_t = \sigma\left(W_f \cdot x_t + W_f \cdot h_{t-1} + b_f\right), \tag{8}$$

where σ is the sigmoid activation function. W_f and b_f are the weight matrix and bias of the forget gate, respectively.

Similarly, the input gate output i_t decides the new input information that should accumulate in the memory cell:

$$i_t = \sigma\left(W_i \cdot x_t + W_i \cdot h_{t-1} + b_i\right), \tag{9}$$

where W_i and b_i are the weight matrix and bias of the input gate, respectively.

The LSTM cell state is then updated as follows, but with conditional self-loop weights W_c and b_c:

$$c_t = f_t \cdot c_{t-1} + i_t \cdot tanh\left(W_c \cdot x_t + W_c \cdot h_{t-1} + b_c\right). \tag{10}$$

The output hidden state h_t of the LSTM cell depends on the cell state c_t and the output gate o_t:

$$h_t = o_t \cdot tanh(c_t). \tag{11}$$

The output gate o_t is calculated using:

$$o_t = \sigma\left(W_o \cdot x_t + W_o \cdot h_t + b_0\right), \tag{12}$$

where W_o and b_o are the weight matrix and bias of the output gate. Note that the hidden state h_t can be shut off via the output gate o_t, which uses a sigmoid activation function.

Encoder-decoder LSTM-RNN further improve the solution quality. This architecture, as depicted in Fig. 5, involves two models: one for reading the input sequence and encoding it into a fixed-length vector, and a second for decoding the fixed-length vector and outputting the predicted sequence. Next section models the PV forecasting problem.

IV. PV FORECASTING

PV forecasting using machine learning involves three steps: (1) Preparing the dataset, (2) Training, and (3) Predicting. Steps one and two are done day-ahead while step three is done in real-time.

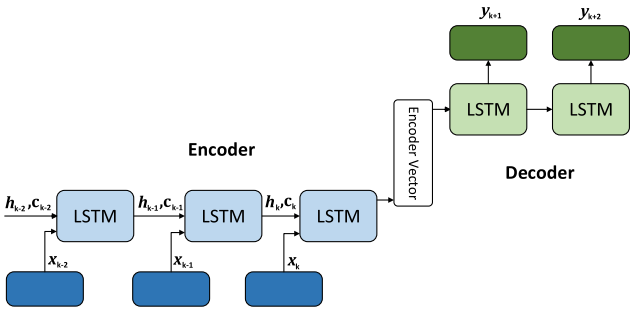


FIGURE 5. Encoder-decoder architecture.

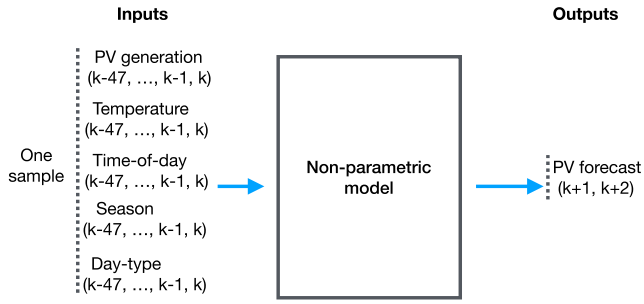


FIGURE 7. Inputs and outputs of the LSTM model.

A. DATA PREPARATION

We used two datasets for the analysis: Dataset A has the actual PV generation from the Arlington Microgrid from June 5, 2019 to June 4, 2020 and Dataset B has the PV generation data generated using the empirical formulas from [31] at the approximate location of the Arlington Microgrid (latitude 48.17 and longitude -122.14) and based on National Renewable Energy Laboratories (NREL) historical solar insolation and meteorological data over 6 years (2013 to 2018).

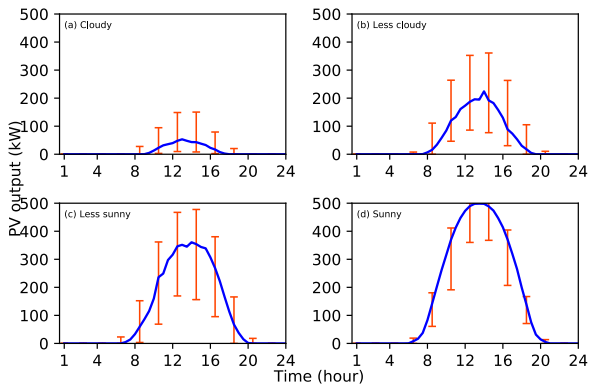


FIGURE 6. Median PV profiles for the four day-types in a year. The upper and lower limits of the error bars show the 95th and 5th quartiles.

Historical data are separated into training and testing sets and arranged in a 3D matrix (samples, time sequence, features) for RNNs and 2D matrix (samples, features) for MLPs. We use PV output, temperature, time-of-the day, season and day-type as features. Days are classified into sunny, less sunny, less cloudy and cloudy day-types by clustering historical PV generation data according to the total daily PV output using a k-means algorithm. Fig. 6 shows the median PV profiles for the four day-types over a year. When predicting, the day-type is estimated on the day ahead using a weather forecasts or a day-ahead PV-forecast. Preliminary simulations showed that the optimum length of the input time-sequence is 47 time-steps (i.e. one day). Outputs are the PV generation for the next two time-steps.

TABLE 1. Simulation parameters of the two datasets.

	Snohomish PUD (∼1 year)	NREL (>6 years)
Features	PV generation (clustered into four day-types)	PV generation Temperature Time-of-day Season Day-type
Training	∼ 2 months for each cluster	5 years
Testing (days)	28 days for each cluster	365 (1 year)
Length of time sequence	One day	One day
LSTM units	25-100	25-100
Epochs	20	20-25
Batch size	1 to 16	1 to 16
Activation	Rectified Linear	Rectified Linear
Optimizer	adam	adam

B. TRAINING

Training the non-parametric model maps inputs and outputs as shown in Fig. 7. The parameters of the LSTM-RNN are given in Table 1.

Since Dataset A covers only a year, we cluster the PV generation data into four day-types according to the total daily PV generation using a k-means algorithm. The aim here is to identify the performance of the machine learning techniques on different types of days. Each of these clusters is trained separately and, because of the limited sample size, only the PV generation was used as a feature (i.e. we implemented a NAR model). On the other hand with the NREL dataset (Dataset B) we were able to use all the features mentioned in Fig. 7 (i.e. we implemented a NARX model). Therefore, we used the Dataset B for the economic analysis.

C. TESTING

The testing dataset was used to make real-time predictions for the next two time-steps, i.e. one hour-ahead forecasts. For Dataset A, 28 days in each cluster are used for testing and the remainder is used for training, as shown in Table 1. On the other hand, the NREL dataset used one year for testing and five years for training.

The machine learning models were trained and tested using Keras, which is a high-level neural networks application programming interface, written in Python, on a computer with a 2.8 GHz Intel Core i7 processor and a 16 GB 2133 MHz LPDDR3 memory card.

V. BATTERY MODELING

Operation of the 1 MW/1.385 MWh Lithium-ion battery of the Arlington Microgrid is limited to the 10% to 90 % SoC range. Fig. 8 shows the expected degradation of this battery over ten years assuming one energy throughput cycle per day, i.e. if there are multiple cycles, the total discharge equals 800 kWh. The average charge and discharge rate is assumed to be 1 C. Note that after ten years, this BESS is expected to have 1 MWh of capacity after degradation.

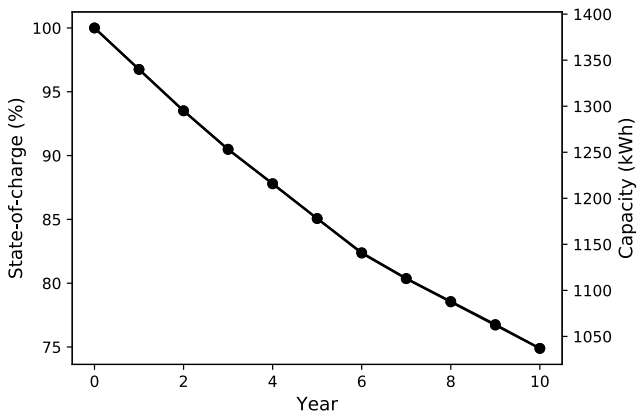


FIGURE 8. Estimated battery life of the Arlington Microgrid's BESS over ten years assuming one energy throughput cycle per day.

In order to compare the battery degradation from the four PV forecasting methods, we generated battery SoC profiles for given actual and forecast PV generation using (1)-(3). The battery charging, battery discharging and inverter efficiencies are assumed to be 98%, 92% and 92%, respectively. The value of C is set to 1.2. Based on these SoC profiles, we use the rainflow algorithm to calculate the battery cycles and their magnitudes.

VI. SIMULATION RESULTS AND DISCUSSION

Figures 9, 10, and 11 show the PV generation forecasts and the battery SoC for typical cloudy, less sunny and sunny days, respectively. Tables 2 (Dataset B) and 3 (Dataset A) compare the RMSE, MAE, MBE and the resulting battery cycles of the LSTM-RNN, encoder-decoder LSTM-RNN, MLPs and the 30/30 persistence forecasting techniques. In all cases, machine learning forecasts are significantly better than the 30/30 persistence forecasts. The RMSE and MAE over a year from the encoder-decoder is 35.7% and 42.6% better than the persistence method (Table 2). The accuracy of the PV forecast varies with the type of day and machine learning technique, as shown in Table 3. For example, LSTM-RNN produced the best PV forecasts on sunny and less sunny days

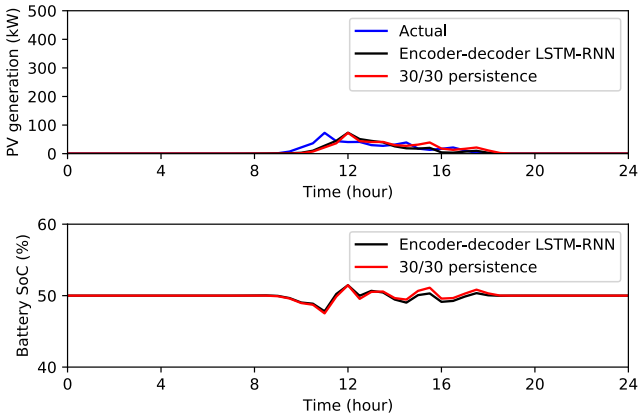


FIGURE 9. PV generation and battery SoC on a cloudy day using encoder-decoder LSTM-RNN and persistence method.

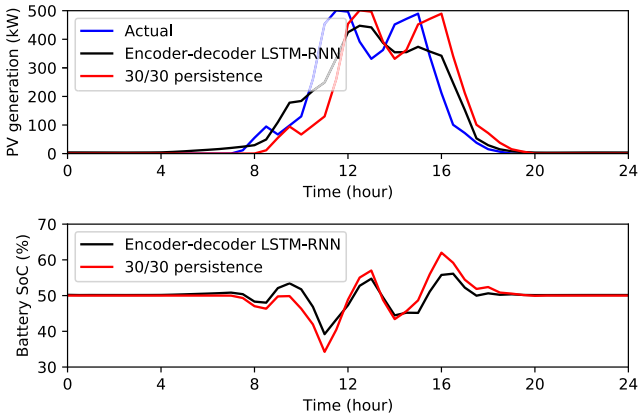


FIGURE 10. PV generation and battery SoC on a less sunny day using encoder-decoder LSTM-RNN and persistence method.

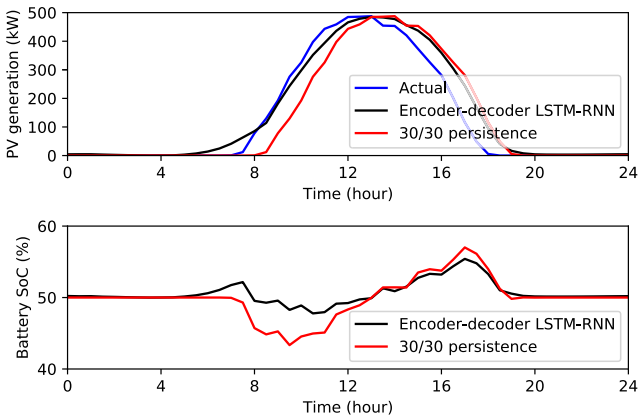


FIGURE 11. PV generation and battery SoC on a sunny day using encoder-decoder LSTM-RNN and persistence method.

while encoder-decoder LSTM-RNN performed the best on less cloudy days.

The number of yearly energy throughput cycles can be reduced by 29.1% (66 cycles per year) using the encoder-decoder LSTM-RNN forecasts. Battery cycles involving more than a 10% Depth of Discharge (DoD) can be reduced

TABLE 2. Simulation results based on NREL data over a year using five years training data. The improvements from the persistence method are given as percentages in parenthesis. The "e", "b" and "h" in the training time column are the epochs, batch size and number of hidden units, respectively. A positive value for MBE indicates an under-prediction while a negative value indicates an over-prediction.

	RMSE (kW)	MAE (kW)	MBE (kW)	Battery cycles (energy throughput)	Battery cycles above 10% DoD	Training time (minutes)
30/30 Persistence	51.12	27.86	0	227	286	-
LSTM-RNN (no day-types)	35.88 (29.8%)	14.41 (48.3%)	0.55	180 (20.7%)	175 (38.8%)	14.7 e=20 b=32 h=100
LSTM-RNN (with day-types)	32.73 (36%)	13.64 (51.0%)	1.24	168 (26%)	153 (46.5%)	14.2 e=20 b=32 h=100
MLP (no day-types)	35.95 (29.7%)	18.18 (34.8%)	-2.1	177 (22%)	163 (43%)	1.3 e=20 b=16 h=200
MLP (with day-types)	34.53 (32.5%)	17.85 (35.9%)	1.9	163 (28.2%)	142 (50.4%)	1.4 e=20 b=16 h=200
Encoder-decoder LSTM-RNN (no day-types)	35.59 (30.4%)	15.79 (43.3%)	1.8	173 (23.8%)	157 (45.1%)	21.7 e=20 b=16 h=50, 25
Encoder-decoder LSTM-RNN (with day-types)	32.87 (35.7%)	15.98 (42.6%)	-1.85	161 (29.1%)	140 (51%)	21.4 e=20 b=16 h=50, 25

by 51%. Such deep cycles have a disproportionate effect on lithium-ion battery degradation. Unfortunately, due to the limited amount of data available about battery degradation, it has not been possible to quantify this effect more accurately. The benefit of using machine learning based forecasts is much higher during sunny days because incorrect forecasts result in higher battery energy throughput, as shown in Fig. 11.

The PV forecast from any of the machine learning techniques are significantly improved by using the type-of-day as a feature, as shown in Table 2. The battery energy throughput, the number of cycles above 10% DoD and the RMSE are improved by 5.3%, 5.9%, and 5.3%, respectively, by using day-type as feature in the encoder-decoder LSTM-RNN. Similar improvements are seen with the MLPs and LSTM-RNN.

In general, over-forecasting and under-forecasting have different effects on the battery usage because the PV output can be curtailed instead of charging the battery to meet the target generation. However, since the PV array of the Arlington Microgrid is part of a community solar project where each panel is owned by a different individual, the PV array is expected to always operate at maximum power (i.e. PV curtailment is not allowed). Given this, the only effect over-forecasting and under-forecasting have in our capacity firming problem is because of the different battery charging and discharging efficiencies. For the sake of completion, we compare the MBE (mean-bias-error) of the PV forecasts

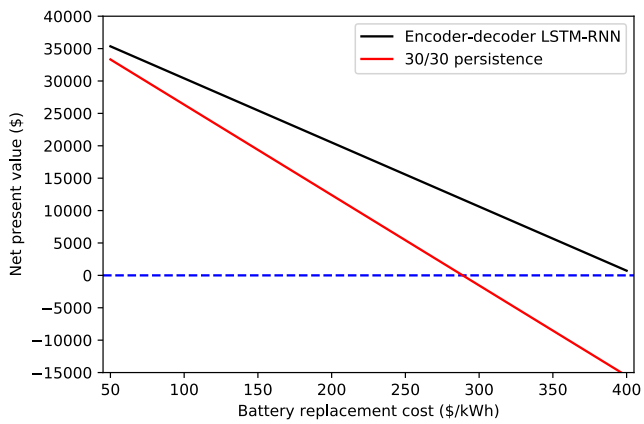


FIGURE 12. Net present value of capacity firming using encoder-decoder LSTM-RNN and 30/30 persistence after 10 years vs. battery replacement cost in 10 years. The revenue is fixed at $5400 per year, the discount rate is assumed to be 5% and the battery degradation per cycle is based on Fig. 8.

in Tables 2 and 3. The small values for MBE in Table 2 means that our yearly forecasts consist of equal amounts of over-forecasts and under-forecasts. According to Table 3, all the machine learning techniques over-forecasts on sunny and less sunny days. This could be because our Dataset A is limited in size.

Tables 2 and 3 show that the time required for training MLPs is significantly smaller than for LSTM-RNN and encoder-decoder LSTM-RNN. However, the testing time is

TABLE 3. Simulation results from actual PV generation data of the Arlington Microgrid that are clustered into four day-types. The improvements from the persistence method are given as percentages in parenthesis. A positive value for MBE indicates an under-prediction while a negative value indicates a over-prediction.

	RMSE (kW)	MAE (kW)	MBE (kW)	Battery cycles (energy throughput)	Battery cycles (above 10% DoD)	Training time (seconds)
Sunny						
30/30 Persistence	68	41.75	0	14.8	30.5	-
LSTM-RNN	32.48 (52.2%)	16.79 (59.8%)	-10.46	9.3 (37.2%)	6 (80.3%)	31
MLP	33.28 (51.1%)	21.86 (47.6%)	-7.72	12.75 (13.9%)	6 (80.3%)	5.7
Encoder-decoder LSTM-RNN	35 (48.6%)	21.8 (47.8%)	-7.42	9.8 (33.8%)	6 (80.3%)	32.1
Less sunny						
30/30 Persistence	70.45	40.76	0	27.9	40.5	-
LSTM-RNN	52. 08 (26.1%)	27.57 (32.4%)	-6.86	22.3 (20.1%)	26 (35.8%)	59
MLP	54.18 (23.1%)	31.43 (22.9%)	-11.87	24.4 (12.5%)	21 (48.1%)	7.2
Encoder-decoder LSTM-RNN	52.2 (25.9%)	31.02 (23.9%)	-12.49	22 (21.1%)	24.5 (39.5%)	55.3
Less cloudy						
30/30 Persistence	63.85	32.99	0	28.3	37.5	-
LSTM-RNN	52.62 (17.6%)	25.95 (21.3%)	4.50	21.7 (23.3%)	25.5 (32%)	43.5
MLP	50.56 (20.8%)	26.85 (18.6%)	-1.35	23.8 (15.9%)	26.5 (29.3%)	6.5
Encoder-decoder LSTM-RNN	49.78 (22%)	24.99 (24.3%)	2.31	21.7 (23.3%)	30 (20%)	42.7
Cloudy						
30/30 Persistence	24.67	10.4	0	8.8	6.5	-
LSTM-RNN	20.93 (15.2%)	8.4 (19.2%)	4.82	7.2 (18%)	3 (53.9%)	71.8
MLP	20.18 (18.2%)	9.24 (11.2%)	2.69	7.8 (11.4%)	2 (69.2%)	7.8
Encoder-decoder LSTM-RNN	21.91 (11.2%)	8.76 (15.8%)	-2.70	7.24 (17.7%)	3 (53.9%)	68.4

similar for all three techniques. Given that the offline training is done on a fast computer, the training time should not be considered when deciding the best PV forecasting method.

Fig. 12 compares how using the best machine learning forecast and persistence forecast affects the NPV (net present value) of capacity firming. Since it is difficult to know what the battery replacement cost will be in ten years, these values have been calculated for a range of replacement costs. This figure shows that encoder-decoder LSTM-RNN based PV forecasting makes capacity firming more profitable than persistence forecasting. If the battery replacement cost is high, capacity firming is not profitable if persistence forecasting is used. In these NPV calculations, the revenue from capacity firming is assumed to be $5400 per year (i.e. the amount that Snohomish PUD would not have to pay to BPA), the interest rate is assumed to be 5%, the battery degradation is assumed to be 0.00687945% per cycle based on Fig 8, and the yearly battery cycles are from Table 2.

Our analysis shows that it is beneficial to do capacity firming if the PV forecasts are based on a state-of-the-art machine learning technique and the capacity firming is already implemented in the BESS or the cost of implementing capacity firming in the BESS is low.

VII. CONCLUSION

This article compared capacity firming using photovoltaic (PV) forecasts based on long short-term memory recurrent neural networks (LSTM-RNN), encoder-decoder LSTM-RNN, multi-layer perceptrons and the 30/30 persistence approaches. The results showed that the encoder-decoder LSTM-RNN performs significantly better than the persistence method in forecasting PV generation and therefore significantly reduces the battery degradation cost of capacity firming.

REFERENCES

[1] *Snohomish County Public Utility District*. Accessed: May 1, 2020. [Online]. Available: https://www.snopud.com/?p=3326

[2] B. Xu, A. Oudalov, A. Ulbig, G. Andersson, and D. Kirschen, "Modeling of lithium-ion battery degradationfor cell life assessment," in *Proc. IEEE Power Energy Soc. Gen. Meeting*, Jul. 2017. p. 1.

[3] B. Xu, J. Zhao, T. Zheng, E. Litvinov, and D. Kirschen, "Factoring the cycle aging cost of batteries participating in electricity markets," in *Proc. IEEE Power Energy Soc. Gen. Meeting (PESGM)*, Aug. 2018. p. 1.

[4] A. Marinopoulos and P. Bakas, "Evaluation of a very large scale PV power system with energy storage for capacity firming," in *Proc. IEEE Eindhoven PowerTech*, Jun. 2015, pp. 1–6.

[5] A. K. Bhattacharjee, I. Batarseh, H. Hu, and N. Kutkut, "An efficient ramp rate and state of charge control for PV-battery system capacity firming," in *Proc. IEEE Energy Convers. Congr. Expo. (ECCE)*, Oct. 2017, pp. 2323–2329.

[6] J. Cardo-Miota, H. Beltran, P. Ayuso, J. Segarra-Tamarit, and E. Perez, "Optimized battery sizing for merchant solar PV capacity firming in different electricity markets," in *Proc. IECON - 45th Annu. Conf. IEEE Ind. Electron. Soc.*, Oct. 2019, pp. 2446–2451.

[7] S. A. Abdelrazek and S. Kamalasadan, "Integrated PV capacity firming and energy time shift battery energy storage management using energy-oriented optimization," *IEEE Trans. Ind. Appl.*, vol. 52, no. 3, pp. 2607–2617, May 2016.

[8] S. Abdelrazek and S. Kamalasadan, "A weather-based optimal storage management algorithm for PV capacity firming," *IEEE Trans. Ind. Appl.*, vol. 52, no. 6, pp. 5175–5184, Nov. 2016.

[9] S. A. Abdelrazek and S. Kamalasadan, "An optimal storage management algorithm for PV capacity firming based on weather patterns," in *Proc. IEEE Ind. Appl. Soc. Annu. Meeting*, Oct. 2015, pp. 1–8.

[10] S. Abdelrazek and S. Kamalasadan, "Integrated control of battery energy storage management system considering PV capacity firming and energy time shift applications," in *Proc. IEEE Ind. Appl. Soc. Annu. Meeting*, Oct. 2014, pp. 1–7.

[11] J. Mathe, N. Miolane, N. Sebastien, and J. Lequeux, "PVNet: A LRCN architecture for spatio-temporal photovoltaic PowerForecasting from numerical weather prediction," 2019, *arXiv:1902.01453*. [Online]. Available: http://arxiv.org/abs/1902.01453

[12] N. Holland, X. Pang, W. Herzberg, S. Karalus, J. Bor, and E. Lorenz, "Solar and PV forecasting for large PV power plants using numerical weather models, satellite data and ground measurements," in *Proc. IEEE 46th Photovoltaic Specialists Conf. (PVSC)*, Jun. 2019, pp. 1609–1614.

[13] M. Husein and I.-Y. Chung, "Day-Ahead Solar Irradiance Forecasting for Microgrids Using a Long Short-Term Memory Recurrent Neural Network: A Deep Learning Approach," *Energies*, vol. 12, no. 10, pp. 1–21, May 2019. [Online]. Available: https://ideas.repec.org/a/gam/jeners/v12y2019i10p1856-d231480.html

[14] Li, Wang, Zhang, Xin, and Liu, "Recurrent neural networks based photovoltaic power forecasting approach," *Energies*, vol. 12, no. 13, p. 2538, Jul. 2019. [Online]. Available: https://www.mdpi.com/1996-1073/12/13/2538

[15] D. Lee and K. Kim, "Recurrent neural network-based hourly prediction of photovoltaic power output using meteorological information," *Energies*, vol. 12, no. 2, p. 215, Jan. 2019. [Online]. Available: https://www.mdpi.com/1996-1073/12/2/215

[16] L. Cho, C.-Y. Hsu, C.-W. Li, W.-C. Chen, H.-R. Wang, and C.-P. Lin, "Study of features and number of time steps to improve hourly solar power forecasting via LSTM," in *Proc. IEEE 8th Global Conf. Consum. Electron. (GCCE)*, Oct. 2019, pp. 303–304.

[17] Y. Yu, J. Cao, and J. Zhu, "An LSTM short-term solar irradiance forecasting under complicated weather conditions," *IEEE Access*, vol. 7, pp. 145651–145666, 2019.

[18] J. Brownlee. (2019). *How to Develop Multi-Step LSTM Time Series Forecasting Models for Power Usage*. [Online]. Available: https://machinelearningmastery.com/how-to-develop-lstm-models-for-multi-step-time-series-forecasting-of-household-power-consumption/

[19] J. Antonanzas, N. Osorio, R. Escobar, R. Urraca, F. J. Martinez-de-Pison, and F. Antonanzas-Torres, "Review of photovoltaic power forecasting," *Sol. Energy*, vol. 136, pp. 78–111, Oct. 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0038092X1630250X

[20] S. Theocharides, R. Alonso-Suarez, G. Giacosa, G. Makrides, M. Theristis, and G. E. Georghiou, "Intra-hour forecasting for a 50 MW photovoltaic system in uruguay: Baseline approach," in *Proc. IEEE 46th Photovoltaic Specialists Conf. (PVSC)*, Jun. 2019, pp. 1632–1636.

[21] Y. Tao and Y. Chen, "Distributed PV power forecasting using genetic algorithm based neural network approach," in *Proc. Int. Conf. Adv. Mech. Syst.*, Aug. 2014, pp. 557–560.

[22] M. Alrashidi, M. Alrashidi, M. Pipattanasomporn, and S. Rahman, "Short-term PV output forecasts with support vector regression optimized by cuckoo search and differential evolution algorithms," in *Proc. IEEE Int. Smart Cities Conf.*, Sep. 2018, pp. 1–8.

[23] G. Mosaico and M. Saviozzi, "A hybrid methodology for the day-ahead PV forecasting exploiting a clear sky model or artificial neural networks," in *Proc. 18th Int. Conf. Smart Technol.*, Jul. 2019, pp. 1–6.

[24] S. Leva, M. Mussetta, A. Nespoli, and E. Ogliari, "PV power forecasting improvement by means of a selective ensemble approach," in *Proc. IEEE Milan PowerTech*, Jun. 2019, pp. 1–5.

[25] F. Wang, Z. Xuan, Z. Zhen, K. Li, T. Wang, and M. Shi, "A day-ahead PV power forecasting method based on LSTM-RNN model and time correlation modification under partial daily pattern prediction framework," *Energy Convers. Manage.*, vol. 212, May 2020, Art. no. 112766. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0196890420303046

[26] Y. Jung, J. Jung, B. Kim, and S. Han, "Long short-term memory recurrent neural network for modeling temporal patterns in long-term power forecasting for solar PV facilities: Case study of South Korea," *J. Cleaner Prod.*, vol. 250, Mar. 2020, Art. no. 119476. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S095965261934346X

[27] M. P. Almeida, M. Muñoz, I. de la Parra, and O. Perpiñán, "Comparative study of PV power forecast using parametric and nonparametric PV models," *Sol. Energy*, vol. 155, pp. 854–866, Oct. 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0038092X17306175

[28] M. Q. Raza, M. Nadarajah, and C. Ekanayake, "On recent advances in PV output power forecast," *Sol. Energy*, vol. 136, pp. 125–144, Oct. 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0038092X16302547

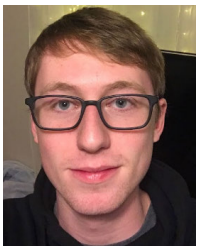
[29] K. Cho, B. van Merrienboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, "Learning phrase representations using RNN encoder-decoder for statistical machine translation," 2014, *arXiv:1406.1078*. [Online]. Available: http://arxiv.org/abs/1406.1078

[30] I. Sutskever, O. Vinyals, and Q. V. Le, "Sequence to sequence learning with neural networks," in *Proc. 27th Int. Conf. Neural Inf. Process. Syst.*, Cambridge, MA, USA, 2014, p. 3104–3112.

[31] M. G. Masters, *Renewable Efficient Electric Power Systerm*. Hoboken, NJ, USA: Wiley, 2004.

CHANAKA KEERTHISINGHE (Member, IEEE) was born in Sri Lanka. He received the B.E. (Hons.) and M.E. degrees in electrical and electronic engineering from The University of Auckland, Auckland, New Zealand, in 2011 and 2014, respectively, and the Ph.D. degree in electrical engineering from The University of Sydney, Sydney, NSW, Australia, in 2017.

He is currently the Lead Electrical Engineer with the Systems Integration Testbed, Washington Clean Energy Testbeds, and a Postdoctoral Researcher with the Department of Electrical and Computer Engineering, University of Washington, Seattle, WA, USA. He has work experience at Mitsubishi Electric Research Laboratories, Cambridge, MA, USA. He holds a U.S. patent and publications in microgrids, energy management in smart homes, wireless power transfer, disaster response, electricity markets, stochastic optimization, and machine learning.



EVAN MICKELSON is currently pursuing the bachelor's degree in electrical engineering with a focus on sustainable power systems with the University of Washington. His mission is to facilitate the integration of clean energy into the grid.

DANIEL S. KIRSCHEN (Fellow, IEEE) received the Ph.D. degree from the University of Wisconsin–Madison and the Electromechanical Engineering degree from the Free University of Brussels, Belgium. He is currently the Donald W. and Ruth Mary Close Professor of electrical engineering with the University of Washington. His research interests include the integration of renewable energy sources in the grid, power system economics, and power system resilience. Prior to joining the University of Washington, he taught for 16 years at The University of Manchester, U.K. Before becoming an Academic, he worked at Control Data Corporation and Siemens on the development of application software for utility control centers. He is the author of two books.

NATHAN SHIH (Student Member, IEEE) is currently pursuing the bachelor's degree in electrical and computer engineering with the University of Washington. He was an Undergraduate Researcher with the University of Washington's Renewable Energy Analysis Laboratory and has previously interned at the National Renewable Energy Laboratory and the NASA Ames Research Center. His research interests include renewable energy, energy poverty, and sustainability.

SCOTT GIBSON (Member, IEEE) received the B.S.E.E. degree from the University of Wyoming, in 1987. He is currently a Principal Engineer with Snohomish County PUD. After graduation, he worked at Boeing and then with a consulting company designing building electrical systems. In 2000, he joined PUD, where he is working on new electrical generation projects. He worked on the development of a tidal generation project and helped to design and constructs two run-of-the-river hydroelectric projects. He is the Project Manager of the PUD's latest energy storage project Arlington Microgrid.